ATTACHMENT VI-2

Statistical Analysis of Groundwater Monitoring Data

1. The Permittee shall perform statistical analysis for Class 1 groundwater monitoring data in accordance with the attached, Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities, Interim final Guidance, April 1989.

2. The Permittee shall enclose a copy of statistical analysis performed in the annual report.

Excerpts from the

STATISTICAL ANALYSIS OF GROUND-WATER MONITORING DATA AT RCRA FACILITIES

INTERIM FINAL GUIDANCE

DRAFT

Office of Solid Waste
Waste Management Division
U.S. Environmental Protection Agency
401 M Street, S.W.
Washington, D.C. 20460

April 1989

7.3 Combined Shewhart-CUSUM Control Charts for Each Well and Constituent

Control charts are widely used as a statistical tool in industry as well as research and development laboratories. The concept of control charts is relatively simple, which makes them attractive to use. From the population distribution of a given variable, such as concentrations of a given constituent, repeated random samples are taken at intervals over time. Statistics, for example the mean of replicate values at a point in time, are computed and plotted together with upper and/or lower predetermined limits on a chart where the x-axis represents time. If a result falls outside these boundaries, then the process is declared to be "out-of-control"; otherwise, the process is declared to be "in control". The widespread use of control charts is due to their ease of construction and the fact that they can provide a quick visual evaluation of a situation, and remedial action can be taken, if necessary.

In the context of ground water monitoring, control charts can be used to monitor the inherent statistical variation of the data collected within a single well, and to flag anomalous results. Further investigation of data points lying outside the established boundaries will be necessary before any direct action is taken.

A control chart that can be used on a real time basis must be constructed from a data set large enough to characterize the behavior of a specific well. It is recommended that data from a minimum of eight samples within a year be collected for each constituent at each well to permit an evaluation of the consistency of monitoring results with the current concept of the hydrogeology of the site. Starks (1988) recommends a minimum of four sampling periods at a unit with eight or more wells and a minimum of eight sampling periods at a unit with less than four wells. Once the control chart for the specific constituent at a given well is acceptable, then subsequent data points can be plotted on it to provide a quick evaluation as to whether the process is in control.

The standard assumptions in the use of control charts are that the data generated by the process, when it is in control, are independently (see Section 2.4.2) and normally distributed with a fixed mean μ and constant variance σ^2. The most important assumption is that of independence; control charts are not robust with respect to departure from independence (e.g., serial correlation, see glossary). In general, the sampling scheme will be such that the possibility of obtaining serially correlated results is minimized, as noted in Section 2. The assumption of normality is of somewhat less concern, but should be investigated before plotting the charts. A transformation (e.g., log-transform, square root transform) can be applied to the raw data so as to obtain errors normally distributed about the mean. An additional situation which may decrease the effectiveness of control charts is seasonality in the data. The problem of seasonality can be

handled by removing the seasonality effect from the data, provided that sufficient data to cover at least two seasons of the same type are available (e.g., 2 years when monthly or quarterly seasonal effect). A procedure to correct a time series for seasonality was shown above in Section 7.2.

PURPOSE

Combined Shewhart-cumulative sum (CUSUM) control charts are constructed for each constituent at each well to provide a visual tool of detecting both trends and abrupt changes in concentration levels.

PROCEDURE

Assume that data from at least eight independent samples of monitoring are available to provide reliable estimates of the mean, μ, and standard deviation, σ, of the constituent's concentration levels in a given well.

Step 1. To construct a combined Shewhart-CUSUM chart, three parameters need to be selected prior to plotting:

h – a decision internal value
k – a reference value
SCL – Shewhart control limit (denoted by U in Starks (1988))

The parameter k of the CUSUM scheme is directly obtained from the value, D, of the displacement that should be quickly detected; k = D/2. It is recommended to select k = 1, which will allow a displacement of two standard deviations to be detected quickly.

When k is selected to be 1, the parameter h is usually set at values of 4 or 5. The parameter h is the value against which the cumulative sum in the CUSUM scheme will be compared. In the context of groundwater monitoring, a value of h = 5 is recommended (Starks, 1988; Lucas, 1982).

The upper Shewhart limit is set at SCL = 4.5 in units of standard deviation. This combination of k = 1, h = 5, and SCL = 4.5 was found most appropriate for the application of combined Shewhart-CUSUM charts for groundwater monitoring (Starks, 1988).

Step 2. Assume that at time period T_i, n_i, concentration measurements X_1, …, X_{ni}, are available. Compute their average X_i.

Step 3. Calculate the standardized mean

$$Z_i = (\overline{X}_i - \mu)\sqrt{n_i} / \sigma$$

where μ and σ are the mean and standard deviation obtained from prior monitoring at the same well (at least four sampling periods in a year).

Step 4. At each time period, T_i, compute the cumulative sum, S_i, as:

$$S_i = \max \{0, (Z_i - k) + S_{i-1}\}$$

where max {A, B} is the maximum of A and B, starting with $S_o = 0$.

Step 5. Plot the values of S_i versus T_i on a time chart for this combined Shewhart-CUSUM scheme. Declare an "out-of-control" situation at sampling period T_i if for the first time, $S_i \geq h$ or $Z_i \geq SCL$. This will indicate probable contamination at the well and further investigations will be necessary.

REFERENCES

Lucas, J. M. 1982. "Combined Shewhart-CUSUM Quality Control Schemes." *Journal of Quality Technology*. Vol. 14, pp. 51-59.

Starks, T. H. 1988 (Draft). "Evaluation of Control Chart Methodologies for RCRA Waste Sites."

Hockman, K. K., and J. M. Lucas. 1987. "Variability Reduction Through Sub-vessel CUSUM Control." *Journal of Quality Technology*. Vol. 19, pp. 113-121.

END OF ATTACHMENT VI-2